

> 0 <
0 | 0 Intelligence
> 0 <
GENERALIGN - Multiple Sequence Alignment Program
Release 5.4

Thu 6 Apr 106 10:05:38-PST

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

4. US-10-756-778-8 (1-492)
1. US-10-756-778-2 (1-742)

Region Alignment: (listed in Clustered order)

US-10-756- 1 mdfsnysseqkypdnnngelicessefyedctnemkkyhpiegdlkfangefdnyyq
consensus mdfsnysseqkypdnnngelicessefyedctnemkkyhpiegdlkfangefdnyyq
US-10-756- 1
US-10-756- 62 hsdvsnaygmktelvtndlpntnminsmntlcrcdlpctmslydnlrstvtvpsfen
consensus hsdvsnaygmktelvtndlpntnminsmntlcrcdlpctmslydnlrstvtvpsfen
US-10-756- 7 TG viTQ
US-10-756- 123 qfdpikflhdieiaietgsfesaITQsmnggrtdiapmlstcfrkvasgallpfpleslgai
consensus qfdpikflhdieiaietgsfesaITQsmnggrtdiapmlstcfrkvasgallpfpleslgai
US-10-756- 13 FRI IndNFI
US-10-756- 184 asfeyrdsqgtgamanlwrqwdvyeRidekil--NFlmgaelaalsakeyarvkvif
consensus asfeyrdsqgtgamanlwrqwdvyeRidekil--NFlmgaelaalsakeyarvkvif
US-10-756- 22 KYIAKLQFSTNOSDLOYPVLTPLRAQACVMHLM
US-10-756- 245 endmntiaepptscgvtlqfllIndnfiKYIAKLQFSTNOSDLOYPVLTPLRAQACVMHLM
consensus endmntiaepptscgvtlqfllIndnfiKYIAKLQFSTNOSDLOYPVLTPLRAQACVMHLM
US-10-756- 56 LLDATTSVWGQOIDSQQLNGYKAEILRLIKVYTNVNTTYNQGLELEKAKPLNYSDBEX
US-10-756- 306 LLDATTSVWGQOIDSQQLNGYKAEILRLIKVYTNVNTTYNQGLELEKAKPLNYSDBEX
consensus LLDATTSVWGQOIDSQQLNGYKAEILRLIKVYTNVNTTYNQGLELEKAKPLNYSDBEX

US-10-756- 117 LQAGRPDISVLSRNFKEVMKNNKVAKYRGMAKMSLSIALFPTRGPNVPRQALKVQSRQ
US-10-756- 367 LQAGRPDISVLSRNFKEVMKNNKVAKYRGMAKMSLSIALFPTRGPNVPRQALKVQSRQ
consensus LQAGRPDISVLSRNFKEVMKNNKVAKYRGMAKMSLSIALFPTRGPNVPRQALKVQSRQ
US-10-756- 178 IFAPVIGIPGIGITSDSGPTFGSMRFVDKTYDQIDALRQMLMELYIOPKSAVFWIYESDMK
US-10-756- 428 IFAPVIGIPGIGITSDSGPTFGSMRFVDKTYDQIDALRQMLMELYIOPKSAVFWIYESDMK
consensus IFAPVIGIPGIGITSDSGPTFGSMRFVDKTYDQIDALRQMLMELYIOPKSAVFWIYESDMK
US-10-756- 239 VRATYVNDYIGKRGSGNTGAAMHMSSDPSALYTSALGAAGYAPNVVGVRYSHGSGYTTGMA
US-10-756- 489 VRATYVNDYIGKRGSGNTGAAMHMSSDPSALYTSALGAAGYAPNVVGVRYSHGSGYTTGMA
consensus VRATYVNDYIGKRGSGNTGAAMHMSSDPSALYTSALGAAGYAPNVVGVRYSHGSGYTTGMA
US-10-756- 300 PANTNAYAPFEFKYPGYTLHSVSAVGLSKAPDADSVMFGRFPVLLNEANQLTDTALQI
US-10-756- 550 PANTNAYAPFEFKYPGYTLHSVSAVGLSKAPDADSVMFGRFPVLLNEANQLTDTALQI
consensus PANTNAYAPFEFKYPGYTLHSVSAVGLSKAPDADSVMFGRFPVLLNEANQLTDTALQI
US-10-756- 361 PAEIGITDVVPAFGRTEBPINGODAIRIMESFTSGFGFTYVDSPOKOKYKIYRIANNLS
US-10-756- 611 PAEIGITDVVPAFGRTEBPINGODAIRIMESFTSGFGFTYVDSPOKOKYKIYRIANNLS
consensus PAEIGITDVVPAFGRTEBPINGODAIRIMESFTSGFGFTYVDSPOKOKYKIYRIANNLS
US-10-756- 422 ASTVSLTYNNQTFPFDILNTSLDPNGVRGNGSYTLVHGPIIIFPSQGTNIFKLASQKEEFA
US-10-756- 672 ASTVSLTYNNQTFPFDILNTSLDPNGVRGNGSYTLVHGPIIIFPSQGTNIFKLASQKEEFA
consensus ASTVSLTYNNQTFPFDILNTSLDPNGVRGNGSYTLVHGPIIIFPSQGTNIFKLASQKEEFA
US-10-756- 483 IDSIIIFSPV
US-10-756- 733 IDSIIIFSPV
consensus IDSIIIFSPV

Alignment score = -768.00

Scoring matrix:

4 1
1 -768

> O <
O| O < Intelligenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Thu 6 Apr 106 10:06:13-PSY

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off

AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

4. US-10-756-778-8 (1-492)
6. US-10-756-778-12 (1-742)

Region Alignment: (listed in Clustered order)

US-10-756- 1 mdpfsnyseqkypdenmgel|txsseqfysdcttenxkxhpieqdl|kfxnqefxdxyq
consensus mdpfsnyseqkypdenmgel|txsseqfysdcttenxkxhpieqdl|kfxnqefxdxyq
US-10-756- 1 hadvsnsgymkteivncl|pynctnxinmwrnt|cxdl|ppecme|ydh|retvrvpsfsn
consensus hadvsnsgymkteivncl|pynctnxinmwrnt|cxdl|ppecme|ydh|retvrvpsfsn
US-10-756- 1 qfcdp|k|f|h|d|e|a|t|g|f|s|a|l|t|c|g|m|n|g|g|c|d|x|m|l|s|t|c|f|f|v|x|a|x|l|p|f|p|s|l|g|a|l
consensus qfcdp|k|f|h|d|e|a|t|g|f|s|a|l|t|c|g|m|n|g|g|c|d|x|m|l|s|t|c|f|f|v|x|a|x|l|p|f|p|s|l|g|a|l
US-10-756- 1 aeflyrvdsqctgaman|wrgmvdvvekr|dsk|l|d|y|n|f|m|g|e|a|a|n|a|s|l|k|e|y|a|r|v|v|k|f|
consensus aeflyrvdsqctgaman|wrgmvdvvekr|dsk|l|d|y|n|f|m|g|e|a|a|n|a|s|l|k|e|y|a|r|v|v|k|f|
US-10-756- 1 IAEPPSTGVITTOFRILINDNFIRIKYIAKLQSTNSDLOYPVLTPLRAQAQCVNHLM
consensus IAEPPSTGVITTOFRILINDNFIRIKYIAKLQSTNSDLOYPVLTPLRAQAQCVNHLM
US-10-756- 245 endmrxAEPPSTGVITTOFRILINDNFIRIKYIAKLQSTNSDLOYPVLTPLRAQAQCVNHLM
consensus endmrxAEPPSTGVITTOFRILINDNFIRIKYIAKLQSTNSDLOYPVLTPLRAQAQCVNHLM
US-10-756- 56 LKDATTSVWGQOIDSQQLNGYKAEILIRLIKVTYNDVNTTNOGLEEKAKPLNTSDPEEX
US-10-756- 306 LKDATTSVWGQOIDSQQLNGYKAEILIRLIKVTYNDVNTTNOGLEEKAKPLNTSDPEEX
consensus LKDATTSVWGQOIDSQQLNGYKAEILIRLIKVTYNDVNTTNOGLEEKAKPLNTSDPEEX

US-10-756- 117 LQAGRPDISVLRNSFKEVMKMKVAKYKGMASALSIALPFTPGPNYPKQALKVYQSRQ
US-10-756- 367 LQAGRPDISVLRNSFKEVMKMKVAKYKGMASALSIALPFTPGPNYPKQALKVYQSRQ
consensus LQAGRPDISVLRNSFKEVMKMKVAKYKGMASALSIALPFTPGPNYPKQALKVYQSRQ

US-10-756- 178 IFAPVIGIPGIGTSODEGPTFGSMRFPDVKTYYDQIDALRGLMELYIOPLKSAFYFYESDMK
US-10-756- 428 IFAPVIGIPGIGTSODEGPTFGSMRFPDVKTYYDQIDALRGLMELYIOPLKSAFYFYESDMK
consensus IFAPVIGIPGIGTSODEGPTFGSMRFPDVKTYYDQIDALRGLMELYIOPLKSAFYFYESDMK

US-10-756- 239 VRATYVNDYIGKRGSTNGAAMWMSDPSAITYSALGAAGYARNVGVARYSHGGSYTKGMA
US-10-756- 489 VRATYVNDYIGKRGSTNGAAMWMSDPSAITYSALGAAGYARNVGVARYSHGGSYTKGMA
consensus VRATYVNDYIGKRGSTNGAAMWMSDPSAITYSALGAAGYARNVGVARYSHGGSYTKGMA

US-10-756- 300 PANTNAYAPPEFKYPGYKLSVSAYGISKAPDAADSVMEGFRPVLLLENANOLLTTALQI
US-10-756- 550 PANTNAYAPPEFKYPGYKLSVSAYGISKAPDAADSVMEGFRPVLLLENANOLLTTALQI
consensus PANTNAYAPPEFKYPGYKLSVSAYGISKAPDAADSVMEGFRPVLLLENANOLLTTALQI

US-10-756- 361 PAEIGITDVVPARGTEEPINGODAIIMESFTSGEGFTYTVDSPOKOKYKIIYRIANNIS
US-10-756- 611 PAEIGITDVVPARGTEEPINGODAIIMESFTSGEGFTYTVDSPOKOKYKIIYRIANNIS
consensus PAEIGITDVVPARGTEEPINGODAIIMESFTSGEGFTYTVDSPOKOKYKIIYRIANNIS

US-10-756- 422 ASTVSLTYNNQTFPTDILNTSLDPNGVGRNGSYTLVEGPIIFESOGTNIIFKLGSQGEFA
US-10-756- 672 ASTVSLTYNNQTFPTDILNTSLDPNGVGRNGSYTLVEGPIIFESOGTNIIFKLGSQGEFA
consensus ASTVSLTYNNQTFPTDILNTSLDPNGVGRNGSYTLVEGPIIFESOGTNIIFKLGSQGEFA

US-10-756- 483 IDSIIIFSPV
US-10-756- 733 IDSIIIFSPV
consensus IDSIIIFSPV

Alignment score = -773.00

Scoring matrix:



> O < Intelligenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Thu 6 Apr 106 10:06:28-PSR

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

4. US-10-756-778-8 (1-492)
7. US-10-756-778-13 (1-492)

Region Alignment: (listed in Clustered order)

US-10-756- 1 IAEPPSTGVITQRIINDNFIKYIAKQFSTNOSDLOQYVLTLPFAQAQCVNHLMLKDAT
US-10-756- 1 xAEPSTGVITQRIINDNFIKYIAKQFSTNOSDLOQYVLTLPFAQAQCVNHLMLKDAT
consensus -AEPSTGVITQRIINDNFIKYIAKQFSTNOSDLOQYVLTLPFAQAQCVNHLMLKDAT
US-10-756- 62 TSVWGGQIDISQQLNGYKAEILIRLIKVTNDVNTTNOGLELEKAKPLANTSDPEEYLQAGR
US-10-756- 62 TSVWGGQIDISQQLNGYKAEILIRLIKVTNDVNTTNOGLELEKAKPLANTSDPEEYLQAGR
consensus TSVWGGQIDISQQLNGYKAEILIRLIKVTNDVNTTNOGLELEKAKPLANTSDPEEYLQAGR
US-10-756- 123 DISVLSNPFKEVKNKVAKYKGMAMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI
US-10-756- 123 DISVLSNPFKEVKNKVAKYKGMAMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI
consensus DISVLSNPFKEVKNKVAKYKGMAMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI
US-10-756- 184 GIPGGITISODSPTFGSMRFVDTYDQIDALRGLMELYIQLKSAVF*IYESDMKVRATYV
US-10-756- 184 GIPGGITISODSPTFGSMRFVDTYDQIDALRGLMELYIQLKSAVF*IYESDMKVRATYV
consensus GIPGGITISODSPTFGSMRFVDTYDQIDALRGLMELYIQLKSAVF*IYESDMKVRATYV
US-10-756- 245 NDYIGRGSNTGAMMWSDDPSA IYTSALGAAGYAPNVGVGRYSHGGSYTKGMAPANTNA
US-10-756- 245 NDYIGRGSNTGAMMWSDDPSA IYTSALGAAGYAPNVGVGRYSHGGSYTKGMAPANTNA
consensus NDYIGRGSNTGAMMWSDDPSA IYTSALGAAGYAPNVGVGRYSHGGSYTKGMAPANTNA
US-10-756- 306 YAPFEKYPGYKLHSYSAVGLSKAPDAASVMEGFRPVLLLENANQLTDTALQIPAEIGI
US-10-756- 306 YAPFEKYPGYKLHSYSAVGLSKAPDAASVMEGFRPVLLLENANQLTDTALQIPAEIGI
consensus YAPFEKYPGYKLHSYSAVGLSKAPDAASVMEGFRPVLLLENANQLTDTALQIPAEIGI

US-10-756- 367 TDVVPARGTEEPINGODAIrIWESFTSGFFTYVDSPQOKXKIIYRIANNLSASTVSL
US-10-756- 367 TDVVPARGTEEPINGODAIxIWESFTSGFFTYVDSPQOKXKIIYRIANNLSASTVSL
consensus TDVVPARGTEEPINGODAIrIWESFTSGFFTYVDSPQOKXKIIYRIANNLSASTVSL

US-10-756- 428 TYNNOFTFDILNTSLDPNGVGRNGYSYTLVEGPIIEFSOGTNIFFKLGSOKGEFALDSIIF
US-10-756- 428 TYNNOFTFDILNTSLDPNGVGRNGYSYTLVEGPIIEFSOGTNIFFKLGSOKGEFALDSIIF
consensus TYNNOFTFDILNTSLDPNGVGRNGYSYTLVEGPIIEFSOGTNIFFKLGSOKGEFALDSIIF

US-10-756- 489 SPV
US-10-756- 489 SPV
consensus SPV

Alignment score = 477.00

Scoring matrix:

4	477
7	7

> O <
o | o < Intelligentics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Thu 6 Apr 106 10:06:39-PST

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off

AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

4. US-10-756-778-8 (1-492)
8. US-10-756-778-14 (1-742)

Region Alignment: (listed in clustered order)

```
US-10-756- 1 mdpfsnyseqkypdnnngelltxsssfydcntnuxkxhpieqdlkfxnqefkdxnyq
Consensus mdpfsnyseqkypdnnngelltxsssfydcntnuxkxhpieqdlkfxnqefkdxnyq

US-10-756- 1
Consensus

US-10-756- 62 hedvsnsgymkteivntdipyntnuxinsmwnlccxdlppecmsiydnlrscvtvpsfsn
Consensus hedvsnsgymkteivntdipyntnuxinsmwnlccxdlppecmsiydnlrscvtvpsfsn

US-10-756- 1
Consensus

US-10-756- 123 qfdqpkfllndieiaixtgsfaalcgennmggtcixpmllstcfffkxaxslpfpjssjgal
Consensus qfdqpkfllndieiaixtgsfaalcgennmggtcixpmllstcfffkxaxslpfpjssjgal

US-10-756- 1
Consensus

US-10-756- 184 asfyvcdsgtgsamanlwrqmwvdyvekrldsklidyhnflimgaelaainaslkeyarvvkif
Consensus asfyvcdsgtgsamanlwrqmwvdyvekrldsklidyhnflimgaelaainaslkeyarvvkif

US-10-756- 1
Consensus

US-10-756- 245 endmnrxaepstgvttofrilndnfrlkiaklqfstnosdlqypvltlplraaqaawhlm
Consensus endmnrxaepstgvttofrilndnfrlkiaklqfstnosdlqypvltlplraaqaawhlm

US-10-756- 56 LKDATTSVWGQOIDSQOLNGYKAEILIRLIKVTYNDVNTTNGLELEKAKPLNYSDEPEY
US-10-756- 306 LKDATTSVWGQOIDSQOLNGYKAEILIRLIKVTYNDVNTTNGLELEKAKPLNYSDEPEY
Consensus LKDATTSVWGQOIDSQOLNGYKAEILIRLIKVTYNDVNTTNGLELEKAKPLNYSDEPEY
```

```
US-10-756- 117 LQAGRPDISVLRSNFKEMKMKVAKYIKRGAMASLSLALFPTPGFNYPKQALKVVOSSHQ
US-10-756- 367 LQAGRPDISVLRSNFKEMKMKVAKYIKRGAMASLSLALFPTPGFNYPKQALKVVOSSHQ
Consensus LQAGRPDISVLRSNFKEMKMKVAKYIKRGAMASLSLALFPTPGFNYPKQALKVVOSSHQ

US-10-756- 178 IFAPVIGIPGIGTSQDSQPTFGSNRPFVKTYYDQIDALRGLMELYIOPLKSAFPIYESDMK
US-10-756- 428 IFAPVIGIPGIGTSQDSQPTFGSNRPFVKTYYDQIDALRGLMELYIOPLKSAFPIYESDMK
Consensus IFAPVIGIPGIGTSQDSQPTFGSNRPFVKTYYDQIDALRGLMELYIOPLKSAFPIYESDMK

US-10-756- 239 VRATYVNDYIGKRGSGNTGAAHWMSSDPSAIYTSALGAAGYAPNVVGVRYSHGGSYTKGMA
US-10-756- 489 VRATYVNDYIGKRGSGNTGAAHWMSSDPSAIYTSALGAAGYAPNVVGVRYSHGGSYTKGMA
Consensus VRATYVNDYIGKRGSGNTGAAHWMSSDPSAIYTSALGAAGYAPNVVGVRYSHGGSYTKGMA

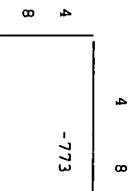
US-10-756- 300 PANTNAYAPPEFKYPGYKLSVSAYGLSKAPDADSVMFGRFPVLLLENANQLLTDTALQI
US-10-756- 550 PANTNAYAPPEFKYPGYKLSVSAYGLSKAPDADSVMFGRFPVLLLENANQLLTDTALQI
Consensus PANTNAYAPPEFKYPGYKLSVSAYGLSKAPDADSVMFGRFPVLLLENANQLLTDTALQI

US-10-756- 361 PAEIGITDVVPAFGRTPEPINGODAIIMESFTSGFGFTYTVDSPOKOKYKIIYRIANNLS
US-10-756- 611 PAEIGITDVVPAFGRTPEPINGODAIIMESFTSGFGFTYTVDSPOKOKYKIIYRIANNLS
Consensus PAEIGITDVVPAFGRTPEPINGODAIIMESFTSGFGFTYTVDSPOKOKYKIIYRIANNLS

US-10-756- 422 ASTVSLTYNNQTFPTDILNTSLDPNGVRGNVGSYTLVEGPIIEFSQGTNIIFKLGSQGEFA
US-10-756- 672 ASTVSLTYNNQTFPTDILNTSLDPNGVRGNVGSYTLVEGPIIEFSQGTNIIFKLGSQGEFA
Consensus ASTVSLTYNNQTFPTDILNTSLDPNGVRGNVGSYTLVEGPIIEFSQGTNIIFKLGSQGEFA

US-10-756- 483 IDSIIIFSPVY
US-10-756- 733 IDSIIIFSPVX
Consensus IDSIIIFSPV-

Alignment score = -773.00
Scoring matrix:
```



> O < IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Thu 6 Apr 106 10:06:52-PST

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

4. US-10-756-778-8 (1-492)
9. US-10-756-778-15 (1-492)

Region Alignment: (listed in clustered order)

```
US-10-756- 1 |AEPSTGVITQRIINDNFIXIALQFSTNQSDIQYVLTLP|P|AQAQCVNHLMLKDAT
US-10-756- 1 |AEPSTGVITQRIINDNFIXIALQFSTNQSDIQYVLTLP|P|AQAQCVNHLMLKDAT
US-10-756- 1 |AEPSTGVITQRIINDNFIXIALQFSTNQSDIQYVLTLP|P|AQAQCVNHLMLKDAT
consensus -AEPSTGVITQRIINDNFIXIALQFSTNQSDIQYVLTLP|P|AQAQCVNHLMLKDAT

US-10-756- 62 |TSVMGQIDISQQLNGYKAEILIRLIKVTYNDVNTYNOGLEKAKPLANSDEEYLQGRP
US-10-756- 62 |TSVMGQIDISQQLNGYKAEILIRLIKVTYNDVNTYNOGLEKAKPLANSDEEYLQGRP
US-10-756- 62 |TSVMGQIDISQQLNGYKAEILIRLIKVTYNDVNTYNOGLEKAKPLANSDEEYLQGRP
consensus |TSVMGQIDISQQLNGYKAEILIRLIKVTYNDVNTYNOGLEKAKPLANSDEEYLQGRP

US-10-756- 123 |DISVLSNPFKEVKKMKVAKYKRGMMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI
US-10-756- 123 |DISVLSNPFKEVKKMKVAKYKRGMMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI
US-10-756- 123 |DISVLSNPFKEVKKMKVAKYKRGMMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI
consensus |DISVLSNPFKEVKKMKVAKYKRGMMASLSIALPPTFGPNYPKQALKVQSRQIFAPVI

US-10-756- 184 |GIPGGITSDQSDPTFGSMRFVDVKTYPQIDALRQIMELYYIQLKSAIFW|YESDMKVRATYV
US-10-756- 184 |GIPGGITSDQSDPTFGSMRFVDVKTYPQIDALRQIMELYYIQLKSAIFW|YESDMKVRATYV
US-10-756- 184 |GIPGGITSDQSDPTFGSMRFVDVKTYPQIDALRQIMELYYIQLKSAIFW|YESDMKVRATYV
consensus |GIPGGITSDQSDPTFGSMRFVDVKTYPQIDALRQIMELYYIQLKSAIFW|YESDMKVRATYV

US-10-756- 245 |NDYIGRGSNTGAMWMSDDPSA|YTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA
US-10-756- 245 |NDYIGRGSNTGAMWMSDDPSA|YTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA
US-10-756- 245 |NDYIGRGSNTGAMWMSDDPSA|YTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA
consensus |NDYIGRGSNTGAMWMSDDPSA|YTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA

US-10-756- 306 |YAPFEFKYPGYKLIHSVAYGLSKAPDADSVNFGFRPVLLLENANQLTDTALQIPAEIGI
US-10-756- 306 |YAPFEFKYPGYKLIHSVAYGLSKAPDADSVNFGFRPVLLLENANQLTDTALQIPAEIGI
US-10-756- 306 |YAPFEFKYPGYKLIHSVAYGLSKAPDADSVNFGFRPVLLLENANQLTDTALQIPAEIGI
consensus |YAPFEFKYPGYKLIHSVAYGLSKAPDADSVNFGFRPVLLLENANQLTDTALQIPAEIGI
```

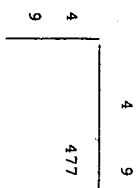
```
US-10-756- 367 |TDVVPARFRTBEPINGODAI|IMESFTSGFPTTYVDSPOKOKYKIIYRIANNLSASTVSL
US-10-756- 367 |TDVVPARFRTBEPINGODAI|IMESFTSGFPTTYVDSPOKOKYKIIYRIANNLSASTVSL
consensus |TDVVPARFRTBEPINGODAI|IMESFTSGFPTTYVDSPOKOKYKIIYRIANNLSASTVSL
```

```
US-10-756- 428 |TYNNQTFPTDILNTSLDPNGVGRNGYSYTLVEGPIIEFSQGTNIFKL|SQKGEPAIDSIIF
US-10-756- 428 |TYNNQTFPTDILNTSLDPNGVGRNGYSYTLVEGPIIEFSQGTNIFKL|SQKGEPAIDSIIF
consensus |TYNNQTFPTDILNTSLDPNGVGRNGYSYTLVEGPIIEFSQGTNIFKL|SQKGEPAIDSIIF
```

```
US-10-756- 489 |SPV
US-10-756- 489 |SPVX
consensus |SPV-
```

Alignment score = 477.00

Scoring matrix:



> O <
O | O Intelligenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Thu 6 Apr 106 10:57:46-PSR

Solution Parameters:

Amino Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

4. US-10-756-778-8 (1-492)
10. US-10-756-778-18 (232-723)

Region Alignment: (listed in Clustered order)

```
US-10-756- 1 |AEPSTGVTORIINDNFIKYIAKLOSTNOSDIQYVLTPLPAQAQCVNHLMLKDAT
US-10-756- 1 |AEPSTGVTORIINDNFIKYIAKLOSTNOSDIQYVLTPLPAQAQCVNHLMLKDAT
consensus -AEPSTGVTORIINDNFIKYIAKLOSTNOSDIQYVLTPLPAQAQCVNHLMLKDAT

US-10-756- 62 |TSVWGQOIDSQOLNGYKAEILIRLIKYYTNDVNTTNOGLEEKAPLNTSDPEEYIQAGR
US-10-756- 62 |TSVWGQOIDSQOLNGYKAEILIRLIKYYTNDVNTTNOGLEEKAPLNTSDPEEYIQAGR
consensus TSVWGQOIDSQOLNGYKAEILIRLIKYYTNDVNTTNOGLEEKAPLNTSDPEEYIQAGR

US-10-756- 123 |DISVLRSNFKFVKWKMKVAKYKRGAMMSALSIALFPFTGPNYPKQALKVQSRQIFAPVI
US-10-756- 123 |DISVLRSNFKFVKWKMKVAKYKRGAMMSALSIALFPFTGPNYPKQALKVQSRQIFAPVI
consensus DISVLRSNFKFVKWKMKVAKYKRGAMMSALSIALFPFTGPNYPKQALKVQSRQIFAPVI

US-10-756- 184 |GIPGGITSDQDSGPTFGSMRFDVKTVDQIDALRLMELYIOPLSAYF*IYESDWKVRATYV
US-10-756- 184 |GIPGGITSDQDSGPTFGSMRFDVKTVDQIDALRLMELYIOPLSAYF*IYESDWKVRATYV
consensus GIPGGITSDQDSGPTFGSMRFDVKTVDQIDALRLMELYIOPLSAYF*IYESDWKVRATYV

US-10-756- 245 |NDYIGRGSNTGAAHMMWSSDPSAIYTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA
US-10-756- 245 |NDYIGRGSNTGAAHMMWSSDPSAIYTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA
consensus NDYIGRGSNTGAAHMMWSSDPSAIYTSALGAAGYAPNVVGVRYSHGGSYTKGMAPANTNA

US-10-756- 306 |YAPFEKYPGYKLHSVSAAYGLSKAPD-ADSVMEGFRPVLLBNANQLTDTALQIPAEIGI
US-10-756- 306 |YAPFEKYPGYKLHSVSAAYGLSKAPD-ADSVMEGFRPVLLBNANQLTDTALQIPAEIGI
consensus YAPFEKYPGYKLHSVSAAYGLSKAPD-ADSVMEGFRPVLLBNANQLTDTALQIPAEIGI
```

```
US-10-756- 367 |TDVVPAFGRTPEEPINGODAI-IMESFTSGFGFTYVDSPOKOKKIIYRIANNLSASTVSL
US-10-756- 367 |TDVVPAFGRTPEEPINGODAI-IMESFTSGFGFTYVDSPOKOKKIIYRIANNLSASTVSL
consensus TDVVPAFGRTPEEPINGODAI-IMESFTSGFGFTYVDSPOKOKKIIYRIANNLSASTVSL
```

```
US-10-756- 428 |TYNNQTFPTDILNTSLDPNGVRGNGSYTLVEGPIIEFSQGTNIFKLGSQKGEPAIDSIIF
US-10-756- 428 |TYNNQTFPTDILNTSLDPNGVRGNGSYTLVEGPIIEFSQGTNIFKLGSQKGEPAIDSIIF
consensus TYNNQTFPTDILNTSLDPNGVRGNGSYTLVEGPIIEFSQGTNIFKLGSQKGEPAIDSIIF
```

```
US-10-756- 489 |SPV
US-10-756- 489 |SPV
consensus SPV-
```

Alignment score = 477.00

Scoring matrix:

4	4	477
10	10	

```

320      330      340      350      360      370      380
FKYPSYKLSHSAVGLSPAPDAADSVGCFRPVLLLENANQLDTALQIAPAEIGITDVVAFPRTEEPING
|||||
FKIPIYKLSHSAVGLSPAPDAADSVGCFRPVLLLENANQLDTALQIAPAEIGITDVVAFPRTEEPING
570      580      590      600      610      620      630
ODAIRIMESFTSGFPTTYSPOKOKKIYRIANNLSASTVSLTNNQTFPTDILNTSLDPNGVGNYS
|||||
ODAIRIMESFTSGFPTTYSPOKOKKIYRIANNLSASTVSLTNNQTFPTDILNTSLDPNGVGNYS
640      650      660      670      680      690      700
460      470      480      490 X
YTLVEGPTEIESOGTNIPIKLSGOKGEFPAIDSIISPPV
|||||
YTLVEGPTEIESOGTNIPIKLSGOKGEFPAIDSIISPPV
710      720      730      740 X

```

2. US-10-756-778-8 (2-492)

US-10-756-778-15 Sequence 15, Application US/10756778

```

Sequence 15, Application US/10756778
GENERAL INFORMATION:
APPLICANT: Agriculture Agroalimentaire Canada
APPLICANT: Cote, Jean-Charles
APPLICANT: Mizuki, Eichi
APPLICANT: Akao, Tetsuyuki
APPLICANT: Jung, Yong-Chul
TITLE OF INVENTION: A NOVEL BACILLUS THURINGIENSIS STRAIN, CRYSTAL GENE AND CRYSTAL
FILE REFERENCE: 12292.5
CURRENT APPLICATION NUMBER: US/10756,778
PRIOR FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: 2,410,153
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.2
SEQ ID NO 15
LENGTH: 492
TYPE: PR
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)..(1)
OTHER INFORMATION: Xaa= isoleucine or methionine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (139)..(139)
OTHER INFORMATION: lysine or arginine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa= serine or histidine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa= glycine or serine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (196)..(196)
OTHER INFORMATION: Xaa= glycine or proline
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (216)..(216)
OTHER INFORMATION: Xaa= glutamine or arginine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (231)..(231)
OTHER INFORMATION: Xaa= tyrosine or tryptophan
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (257)..(257)

```

```

OTHER INFORMATION: Xaa= alanine or leucine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (260)..(260)
OTHER INFORMATION: Xaa= glycine or histidine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (268)..(268)
OTHER INFORMATION: Xaa= alanine or valine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (301)..(301)
OTHER INFORMATION: Xaa= alanine or proline
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (387)..(387)
OTHER INFORMATION: Xaa= arginine or isoleucine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (475)..(475)
OTHER INFORMATION: Xaa= glycine or arginine
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (492)..(492)
OTHER INFORMATION: Xaa= valine or serine

```

Initial Score = 477 Optimized Score = 477 Significance = -0.33
 Residue Identity = 96% Matches = 477 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

```

X      10      20      30      40      50      60      70
IAPPESTGYTQFRINDNFIRKIAKQSTNSQSLQYVPLTLPLRAQCVNHLMLKQATTSVWGQOIDSQ
|||||
XAPPESTGYTQFRINDNFIRKIAKQSTNSQSLQYVPLTLPLRAQCVNHLMLKQATTSVWGQOIDSQ
10      20      30      40      50      60      70

```

```

QUNGYKAEILRLIKVTYNDVNTTNGLELEKAKPLNYSDEPEYLQAGPDISVLRSNKFEVWKNKVAKKY
|||||
QUNGYKAEILRLIKVTYNDVNTTNGLELEKAKPLNYSDEPEYLQAGPDISVLRSNKFEVWKNKVAKKY
80      90      100      110      120      130      140

```

```

RGMAWSALSLAALFPTFGSNYPKQALKVQSRQIFAPVIGIGITTSQDSGPTFGSMRPDVXTYQIDALRQ
|||||
RGMAWSALSLAALFPTFGSNYPKQALKVQSRQIFAPVIGIGITTSQDSGPTFGSMRPDVXTYQIDALRQ
150      160      170      180      190      200      210

```

```

LMELIYIOPLSAIFWIIYESDWRATYVNDYIGKGSNTGAAMHWSDDPSAIYTSALGAAYAPNVGVRY
|||||
LMELIYIOPLSAIFWIIYESDWRATYVNDYIGKGSNTGAAMHWSDDPSAIYTSALGAAYAPNVGVRY
220      230      240      250      260      270      280

```

```

SHGGSYTKGAPANTNAVAPPEFKIPGYLHSVSAVGLSKAPDAADSVGFRPVLLLENANQLDTALQI
|||||
SHGGSYTKGAPANTNAVAPPEFKIPGYLHSVSAVGLSKAPDAADSVGFRPVLLLENANQLDTALQI
290      300      310      320      330      340      350      360

```

```

PAEIGITDVVPAFGRTEEPINGODAIRIMESFTSGFPTTYSPOKOKKIYRIANNLSASTVSLTNNQ
|||||
PAEIGITDVVPAFGRTEEPINGODAIRIMESFTSGFPTTYSPOKOKKIYRIANNLSASTVSLTNNQ
370      380      390      400      410      420      430

```

```

TFPTDILNTSLDPNGVGNYSYTLVEGPTEIESOGTNIPIKLSGOKGEFPAIDSIISPPV
|||||
TFPTDILNTSLDPNGVGNYSYTLVEGPTEIESOGTNIPIKLSGOKGEFPAIDSIISPPV
440      450      460      470      480      490 X

```

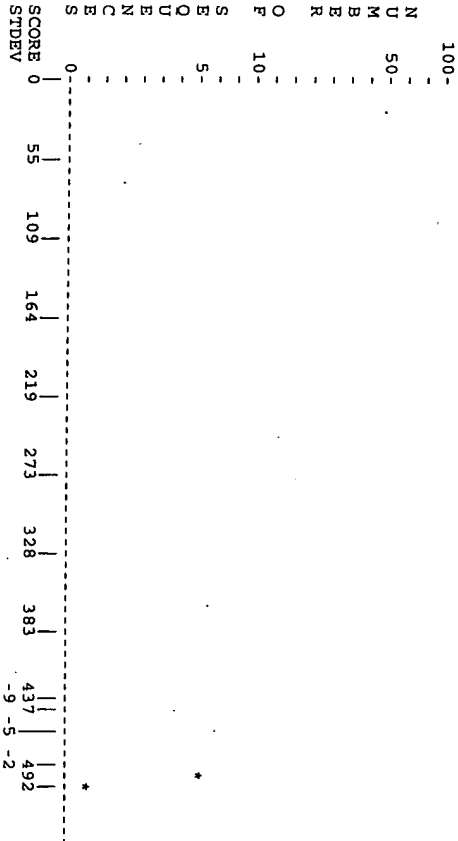

> O <
O | 10 Intelligence
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-10-756-778-8.res made by bshears on Thu 6 Apr 106 10:09:35-PST.

Query sequence being compared: US-10-756-778-8 (1-492)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of US-10-756-778-8 (1-492) with:
File: /home/bshears/rook*pep



PARAMETERS

Similarity matrix Unitary 1 K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 32
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 479 Median 478 Standard Deviation 6.12
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 3933
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Score	Init. Opt.	Sig. Frame
---------------	-------------	--------	-------	------------	------------

The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig. Frame
1. US-10-756-778-2	Sequence 2, Application US	742	492	492	2.12 0

Sequence Name	Description	Length	Score	Init. Opt.	Sig. Frame
2. US-10-756-778-15	Sequence 15, Application	492	477	477	-0.33 0
3. US-10-756-778-13	Sequence 13, Application	492	477	477	-0.33 0
4. US-10-756-778-18	Sequence 18, Application	723	477	477	-0.33 0
5. US-10-756-778-14	Sequence 14, Application	742	477	477	-0.33 0
6. US-10-756-778-12	Sequence 12, Application	742	477	477	-0.33 0

1. US-10-756-778-8 (1-492)
US-10-756-778-2 Sequence 2, Application US/10756778

Sequence 2, Application US/10756778

GENERAL INFORMATION:
APPLICANT: Agriculture Agrolimentaire Canada
APPLICANT: Cote, Jean-Charles
APPLICANT: Mizuki, Eiichi
APPLICANT: Akao, Tetsuyuki
APPLICANT: Jung, Yong-Chul
TITLE OF INVENTION: A NOVEL BACILLUS THURINGIENSIS STRAIN, CRYSTAL GENE AND CRYSTAL
FILE REFERENCE: 12292.5
CURRENT APPLICATION NUMBER: US/10/756, 778
CURRENT FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: 2,410,153
PRIOR FILING DATE: 2002-12-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent version 3.2
SEQ ID NO 2
LENGTH: 742
TYPE: PRT
ORGANISM: Bacillus thuringiensis

Initial Score = 492 Optimized Score = 492 Significance = 2.12
Residue identity = 100% Matches = 492 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

Sequence	US-10-756-778-8	US-10-756-778-2
1	RMWDVYERKIDSKILDYHNFIMGAELALNASTLKEVAVVYKIFENDMNRIAEPSTGVITQFRILINDFIK	IAEPSTGVITQFRILINDFIK
2	210 220 230 240 250 260 270	
3	YIAKLOFSTNGSDLOYPVLTLPPLRAQACVMMMLKDKATTSVWGQIDISQQLNGYKAEIRIRIKYITNDVNT	
4	280 290 300 310 320 330 340	
5	YIAKLOFSTNGSDLOYPVLTLPPLRAQACVMMMLKDKATTSVWGQIDISQQLNGYKAEIRIRIKYITNDVNT	
6	350 360 370 380 390 400 410	
7	YIAKLOFSTNGSDLOYPVLTLPPLRAQACVMMMLKDKATTSVWGQIDISQQLNGYKAEIRIRIKYITNDVNT	
8	420 430 440 450 460 470 480	
9	YIAKLOFSTNGSDLOYPVLTLPPLRAQACVMMMLKDKATTSVWGQIDISQQLNGYKAEIRIRIKYITNDVNT	
10	490 500 510 520 530 540 550 560	

PRIOR APPLICATION NUMBER: 2,410,153
 PRIOR FILING DATE: 2002-12-05
 NUMBER OF SEQ.ID NOS: 18
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 18
 LENGTH: 723
 TYPE: PRF
 ORGANISM: Bacillus thuringiensis

Initial Score = 477 Optimized Score = 477 Significance = -0.33
 Residue Identity = 96% Matches = 477 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

ROMVDYERKIDSKLDYHNFIMGAELALNALKEVARVVKIFENDNRRAAEPPSTGVITQFRLINDNFIX
 190 200 210 220 230 X 240 250
 30 40 50 60 70 80 90
 YIAAQFSTNSDLYPVLTLPLRAQACVMHMLKIDATTSVMGQIDISQOLNGYKALRLIKVYNDVNT
 260 270 280 290 300 310 320
 YIALQFSTNSDLYPVLTLPLRAQACVMHMLKIDATTSVMGQIDISQOLNGYKALRLIKVYNDVNT
 100 110 120 130 140 150 160
 TYNGLELEKAKPLNYSPEBYLQAGRDIVLRSNFEVKMKNVAKYKKGMMASLSIALPPTGPNYP
 330 340 350 360 370 380 390
 TYNGLELEKAKPLNYSPEBYLQAGRDIVLRSNFEVKMKNVAKYKKGMMASLSIALPPTGPNYP
 170 180 190 200 210 220 230
 KOAKKVOSRQIFAPVIGIRGITSQDSGPTFGSRFVKTIDIDALRLMELYIOLKSAFYIYESDWK
 400 410 420 430 440 450 460
 KOAKKVOSRQIFAPVIGIRGITSQDSGPTFGSRFVKTIDIDALRLMELYIOLKSAFYIYESDWK
 240 250 260 270 280 290 300 310
 VRATYVDYIKRGSTNGAAMHSSSDPSATITSAAGAGYAPNVGVYSHGSGYTKGMAPNTNAYAPE
 470 480 490 500 510 520 530 540
 VRATYVDYIKRGSTNGAAMHSSSDPSATITSAAGAGYAPNVGVYSHGSGYTKGMAPNTNAYAPE
 320 330 340 350 360 370 380
 FKYPGYKLHSAVGLSAPPAOSVMGFPRVLIENANOLITDTALQIPAEIGITDVVAFKRTPEPING
 550 560 570 580 590 600 610
 FKYPGYKLHSAVGLSAPPAOSVMGFPRVLIENANOLITDTALQIPAEIGITDVVAFKRTPEPING
 390 400 410 420 430 440 450
 ODAIRIMESFTSGFGFTYTVDSPOKOKKIIYRIANNISASTVSLTYNNQTFPDILNTSLDPNGVGNYS
 620 630 640 650 660 670 680
 ODAIRIMESFTSGFGFTYTVDSPOKOKKIIYRIANNISASTVSLTYNNQTFPDILNTSLDPNGVGNYS
 460 470 480 490 X
 YTVVGPPIESGGINIFKLGSQKGEPAIDSIIFSPV
 690 700 710 720 X
 YTVVGPPIESGGINIFKLGSQKGEPAIDSIIFSPV

5. US-10-756-778-8 (1.492)
 US-10-756-778-14 Sequence 14, Application US/10756778

Sequence 14, Application US/10756778
 GENERAL INFORMATION:
 APPLICANT: Agriculture Agroalimentaire Canada
 APPLICANT: Cote, Jean-Charles
 APPLICANT: Mizuki, Eiichi
 APPLICANT: Akao, Tetsuyuki
 APPLICANT: Jung, Yong-Chul
 TITLE OF INVENTION: A NOVEL BACILLUS THURINGIENSIS STRAIN, CRYSTAL GENE AND CRYSTAL

TITLE OF INVENTION: PROTEIN AND USES THEREOF
 FILE REFERENCE: 12292.5
 CURRENT APPLICATION NUMBER: US/10/756, 778
 CURRENT FILING DATE: 2004-01-14
 PRIOR APPLICATION NUMBER: 2,410,153
 PRIOR FILING DATE: 2002-12-05
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 14
 LENGTH: 742
 TYPE: PRF
 ORGANISM: Bacillus thuringiensis

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (24)..(24)
 OTHER INFORMATION: Xaa=glutamate or lysine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (37)..(37)
 OTHER INFORMATION: Xaa= methionine or alanine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (39)..(39)
 OTHER INFORMATION: Xaa= threonine or asparagine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (51)..(51)
 OTHER INFORMATION: Xaa= alanine or threonine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (56)..(56)
 OTHER INFORMATION: Xaa= proline or serine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (59)..(59)
 OTHER INFORMATION: Xaa= tyrosine or tryptophan
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (87)..(87)
 OTHER INFORMATION: Xaa= asparagine or aspartate
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (97)..(97)
 OTHER INFORMATION: Xaa= arginine or lysine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (138)..(138)
 OTHER INFORMATION: Xaa= glutamate or lysine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (158)..(158)
 OTHER INFORMATION: Xaa= alanine or asparagine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (170)..(170)
 OTHER INFORMATION: Xaa= glycine or serine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (251)..(251)
 OTHER INFORMATION: Xaa= isoleucine or methionine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (389)..(389)
 OTHER INFORMATION: Xaa= lysine or arginine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (444)..(444)
 OTHER INFORMATION: Xaa= serine or histidine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (445)..(445)
 OTHER INFORMATION: Xaa= glycine or serine
 FEATURE:

NAME/KEY: MISC FEATURE
 LOCATION: (446)..(446)
 OTHER INFORMATION: Xaa= glycine or proline
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (466)..(466)
 OTHER INFORMATION: Xaa= glutamine or arginine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (481)..(481)
 OTHER INFORMATION: Xaa= tyrosine or tryptophan
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (507)..(507)
 OTHER INFORMATION: Xaa= alanine or leucine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (510)..(510)
 OTHER INFORMATION: Xaa= glycine or histidine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (518)..(518)
 OTHER INFORMATION: Xaa= alanine or valine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (551)..(551)
 OTHER INFORMATION: Xaa= alanine or proline
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (582)..(582)
 OTHER INFORMATION: Xaa= alanine or threonine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (637)..(637)
 OTHER INFORMATION: Xaa= arginine or isoleucine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (725)..(725)
 OTHER INFORMATION: Xaa= glycine or arginine
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (742)..(742)
 OTHER INFORMATION: Xaa= valine or serine

Initial Score = 477 Optimized Score = 477 Significance = -0.33
 Residue Identity = 96% Matches = 477 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

RQWVDYERKIDSKILDYHNFIMGAEALALNASLKEYARVVKIPENIMNRXAEPPSTGVITOFRIINDNFIK
 210 220 230 240 250 260 270

YIAKLGSTNOSDIQYVLTLPFAQACVWHLMLKDAITTSVWGQOIDSQGLNGYKELIRLIKIVYNDVT
 30 40 50 60 70 80 90
 YIAKLGSTNOSDIQYVLTLPFAQACVWHLMLKDAITTSVWGQOIDSQGLNGYKELIRLIKIVYNDVT
 280 290 300 310 320 330 340

TYNGLLEKAKPLNYSDDPEEYLDAGRPDISVLSNFEVWKNKVKAKYRGWMSALSLAALFPTGPNYP
 100 110 120 130 140 150 160
 TYNGLLEKAKPLNYSDDPEEYLDAGRPDISVLSNFEVWKNKVKAKYRGWMSALSLAALFPTGPNYP
 350 360 370 380 390 400 410

KQALKVQSRQIFAPVIGIGITTSQDSGPTFGSMRPDVKTYYDIDALROLMEYIOLKSAFYWIESDMK
 170 180 190 200 210 220 230
 KQALKVQSRQIFAPVIGIGITTSQDSGPTFGSMRPDVKTYYDIDALROLMEYIOLKSAFYWIESDMK
 420 430 440 450 460 470 480

VBATVNDYIGKRGSGNTGAAMHMSDDPSAIYTSALGAAGYAPNVGVRYHSGSYTKGMAPNTNAYAPFE
 490 500 510 520 530 540 550 560
 VBATVNDYIGKRGSGNTGAAMHMSDDPSAIYTSALGAAGYAPNVGVRYHSGSYTKGMAPNTNAYAPFE
 320 330 340 350 360 370 380
 FKYPGKHSVSAVYGSAPDADSVSMGFRVLLNEANOLTTTALQIPREIGITTVVPAFGTEERING
 570 580 590 600 610 620 630
 FKYPGKHSVSAVYGSAPDADSVSMGFRVLLNEANOLTTTALQIPREIGITTVVPAFGTEERING
 390 400 410 420 430 440 450
 QDAIRWESFTSGPGETTVDSPOKOKIIRIANNLSASVSTIYNNQFFIILNTSLDPNVRGNGYS
 640 650 660 670 680 690 700
 QDAIRWESFTSGPGETTVDSPOKOKIIRIANNLSASVSTIYNNQFFIILNTSLDPNVRGNGYS
 460 470 480 490 X
 YTLVSGPIIEFSQGTNIFPLSGOKGEPAIDSIIEFPV
 710 720 730 740 X
 YTLVSGPIIEFSQGTNIFPLSGOKGEPAIDSIIEFPV

6. US-10-756-778-8 (1-492) US-10-756-778-12 Sequence 12, Application US/10756778

Sequence 12, Application US/10756778
 GENERAL INFORMATION:

APPLICANT: Agriculture Agralimmentaire Canada
 APPLICANT: Cote, Jean-Charles

APPLICANT: Mizuki, Ichichi

APPLICANT: Akao, Tetsuyuki

APPLICANT: Jung, Yong-Chul

TITLE OF INVENTION: A NOVEL BACILLUS THURINGIENSIS STRAIN, CRYSTAL GENE AND CRYSTAL

FILE REFERENCE: 12292.5

CURRENT APPLICATION NUMBER: US/10/756,778

CURRENT FILING DATE: 2004-01-14

PRIOR APPLICATION NUMBER: 2,410,153

PRIOR FILING DATE: 2002-12-05

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12

LENGTH: 742

TYPE: PR

ORGANISM: Bacillus thuringiensis

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (24)..(24)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (37)..(37)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (39)..(39)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (51)..(51)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (56)..(56)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (59)..(59)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (59)..(59)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (59)..(59)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (59)..(59)

OTHER INFORMATION: Xaa=any amino acid

FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (59)..(59)

OTHER INFORMATION: Xaa=any amino acid

LOCATION: (87)..(87)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (97)..(97)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (138)..(138)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (158)..(158)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (170)..(170)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (251)..(251)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (389)..(389)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (444)..(444)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (445)..(445)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (466)..(466)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (467)..(467)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (481)..(481)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (507)..(507)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (510)..(510)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (518)..(518)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (551)..(551)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (582)..(582)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (637)..(637)
OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (725)..(725)

OTHER INFORMATION: Xaa=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (742)..(742)
OTHER INFORMATION: Xaa=any amino acid

Initial Score = 477 Optimized Score = 477 Significance = -0.33
Residue Identity = 96% Matches = 477 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

QDAIRIWESFTSGFGFTTYVDSPQKOKYIIRIANNLASVLSLTYNNQTFETDILNTSLDPNGVGRNGYS
QDAIXIWESFTSGFGFTTYVDSPQKOKYIIRIANNLASVLSLTYNNQTFETDILNTSLDPNGVGRNGYS
460 470 480 490 X
YTLVSGPIIEFSGQTNIFPKLGSQKGEFAIDSIIFSPV
YTLVSGPIIEFSGQTNIFPKLGSQKGEFAIDSIIFSPV
710 720 730 740 X

VRATVNDYIGKRGNTGAMHMMSSDPXIIYTSALGAGVAPNVGVYSHSGSYTKGMADPNTNAYAPFE
VRATVNDYIGKRGNTGAMHMMSSDPXIIYTSALGAGVAPNVGVYSHSGSYTKGMADPNTNAYAPFE
490 500 510 520 530 540 550 560

FKYPGYKLHVSAYGLSKAPDADSVMFGEFVLLNENANQLTDTALQIPAIIGITDVPAPGRTPEEPNG
FKYPGYKLHVSAYGLSKAPDADSVMFGEFVLLNENANQLTDTALQIPAIIGITDVPAPGRTPEEPNG
570 580 590 600 610 620 630

YIAKLOPSTNOSDLOYPVLTPLPRAQACVHMLMLKDATTTSWGQOIDSQOLNGYAEILRLIKVYTNVNT
YIAKLOPSTNOSDLOYPVLTPLPRAQACVHMLMLKDATTTSWGQOIDSQOLNGYAEILRLIKVYTNVNT
280 290 300 310 320 330 340

TYNGLLEKAKPLINYSDBPEYLOAGRDPISVLRSNFKEVMKNKVAKTKRGAMASLALPFPFGPNYP
TYNGLLEKAKPLINYSDBPEYLOAGRDPISVLRSNFKEVMKNKVAKTKRGAMASLALPFPFGPNYP
350 360 370 380 390 400 410

KOALKVQSRQIFAPVIGLPGGITSDQXXXTFGSMFVKYTDQIDALXLMELIYQPLKSAVFXIYESDWK
KOALKVQSRQIFAPVIGLPGGITSDQXXXTFGSMFVKYTDQIDALXLMELIYQPLKSAVFXIYESDWK
420 430 440 450 460 470 480

VRATVNDYIGKRGNTGAMHMMSSDPXIIYTSALGAGVAPNVGVYSHSGSYTKGMADPNTNAYAPFE
VRATVNDYIGKRGNTGAMHMMSSDPXIIYTSALGAGVAPNVGVYSHSGSYTKGMADPNTNAYAPFE
240 250 260 270 280 290 300 310

> O <
O| |O Intelligence
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-756-778-8.res made by behears on Thu 6 Apr 106 10:55:54-PST.

Query sequence being compared: US-10-756-778-8 (1-492)
Number of sequences searched: 1
Number of scores above cutoff: 1

Results of the initial comparison of US-10-756-778-8 (1-492) with:
File: /home/behears/rook756778b.pap

```

100-
N -
U 50-
M -
B -
E -
R -
O 10-
S -
E 5-
O -
U -
N -
C -
E -
S -
SCORE 0 53 106 159 212 265 318 371 424 477
SIDEV

```

PARAMETERS

Similarity matrix Unitary 1 K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 32
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 477 Median 0 Standard Deviation 0.00
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 492
Number of sequences searched: 1
Number of scores above cutoff: 1

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt. Frame	Sig.
1. US-10-756-778-18	Sequence 18, Application	492	477	477	0.00

1. US-10-756-778-8 (1-492) → Residues 232-723

US-10-756-778-18 Sequence 18, Application US/10756778

Sequence 18, Application US/10756778
GENERAL INFORMATION:

APPLICANT: Agriculture Agroalimentaire Canada
APPLICANT: Cote, Jean-Charles
APPLICANT: Mizuki, Eichi
APPLICANT: Akao, Tetsuyuki
TITLE OF INVENTION: A NOVEL BACILLUS THURINGIENSIS STRAIN, CRYSTAL GENE AND CRYSTAL
FILE REFERENCE: 12292.5
CURRENT APPLICATION NUMBER: US/10/756,778
CURRENT FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: 2,410,153
PRIOR FILING DATE: 2002-12-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent version 3.2
SEQ ID NO: 18
LENGTH: 723
TYPE: PRT
ORGANISM: Bacillus thuringiensis

Initial Score = 477 Optimized Score = 477 Significance = 0.00
Residue Identity = 96% Matches = 477 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

X 10 20 30 40 50 60 70
IAEPSTGVITOFRIINDNFIKYIAKLOFSTNOSDLOYVLTLPRAQACVHMLLKDATTSVWGQIDSO
MAEPSTGVITOFRIINDNFIKYIAKLOFSTNOSDLOYVLTLPRAQACVHMLLKDATTSVWGQIDSO
X 10 20 30 40 50 60 70

```

```

80 90 100 110 120 130 140
OLNGYKALIRLIKVTYNDVNTYNOGLEAKPLNYSDEEYIOARPDIVLSNPKVMMKRVAKYK
OLNGYKALIRLIKVTYNDVNTYNOGLEAKPLNYSDEEYIOARPDIVLSNPKVMMKRVAKYK
80 90 100 110 120 130 140

```

```

150 160 170 180 190 200 210
RGMASALSLALPFTFGPNYPKQALKVOSROIFAPVIGIPGGITSDSGPTGSMRFDVKTVDIDALRQ
RGMASALSLALPFTFGPNYPKQALKVOSROIFAPVIGIPGGITSDSGPTGSMRFDVKTVDIDALRQ
150 160 170 180 190 200 210

```

```

220 230 240 250 260 270 280
IMELYIOPKSAFYFWISDMKVRATYVNDYIGKSGNTGAAMHWSDEPAITYSALGAGYAPNVGVRY
IMELYIOPKSAFYFWISDMKVRATYVNDYIGKSGNTGAAMHWSDEPAITYSALGAGYAPNVGVRY
220 230 240 250 260 270 280

```

```

290 300 310 320 330 340 350
SHGSGYTKGMAMPNTNNAVAPEPFKPGYKLSVSAYGLSKAPDAADSVMPFRPVLLENENQLLTDLAI
SHGSGYTKGMAMPNTNNAVAPEPFKPGYKLSVSAYGLSKAPDAADSVMPFRPVLLENENQLLTDLAI
290 300 310 320 330 340 350

```

```

370 380 390 400 410 420 430
PAEIGTIVPAFGRTSPINGODAIRIWESEFTSGFTYVDSPOKQYKIVRIANLSASTVSLTNNQ
PAEIGTIVPAFGRTSPINGODAIRIWESEFTSGFTYVDSPOKQYKIVRIANLSASTVSLTNNQ
370 380 390 400 410 420 430

```

```

440 450 460 470 480 490 X
TFEFDILNLSLDPNGVRNGYGYTLVVEGPIIEFGSGTNIIFKLGSGKGFALDISITFSRV
TFEFDILNLSLDPNGVRNGYGYTLVVEGPIIEFGSGTNIIFKLGSGKGFALDISITFSRV
440 450 460 470 480 490 X

```

Thu Apr 6 10:58:44 2006

us-756-778-8.res

Page 2

TFPTDINTSLDNGVGRNGVSYTLVESPILIESQCTNIFKLRSQKGEPAIDSIIFSPVS
440 450 460 470 480 490 X
